A Cognitive Load Approach To Learner-Centered Design Of Digital Instructional Media And Supporting Accessibility Tools

Ashley Russell  Daniel Hannon
Cambridge Consultants  Tufts University

Instructional material provided through digital media and various types of educational technology may be able to support students with a wide range of learning abilities, providing better educational opportunities for a greater number of students. Technology-based instructional design, however, often does not take into consideration variation in student abilities. Research on Universal Design for Learning (UDL) promotes the idea of providing students with a variety of means of representation and expression. Investigations into cognitive load note the impact of the complexity of information content (intrinsic cognitive load) and the type of information presentation (extraneous cognitive load) on student learning. The work provided here is part of a larger effort to understand the influence of a student-centered approach to technology-based, instructional design in the field of high school chemistry. In the present study, results from a study by Carlson, Chandler, and Sweller (2003) of cognitive load were replicated in a digital media learning environment to determine whether the elements of cognitive load theory can be extended into design requirements for technology-based instructional materials. In a two-by-two experimental design, participants were required to learn contents in organic chemistry that was designed with either high or low intrinsic load and high or low extraneous load. Results were congruent with Carlson et al. (2003) and further support the following claims: (1) when compared to low element interactivity, high intrinsic load leads to decreased performance and higher subjective workload, (2) extraneous load reduction can promote performance gains and reduction in workload for tasks with high element interactivity.

I. Introduction & Background

In the digital age, technology is “increasingly a means for empowering students, a method for communication and socializing, and a ubiquitous, transparent part of their lives” (Johnson, Smith, Levine, & Haywood, 2010). The recent changes in technology, the adoption of technology into the daily lives of students, and lagging educational systems contribute to an overarching need for computer assimilation in the classroom. Additionally, recent changes in legislation promote an integrated student body, one that includes students with disabilities, which further endorse the need for tools that allow all learners access to information at their own pace and subsequently provide a successful learning environment (Rose, Meyer, & Hitchcock, 2005). Educational technology could be such a tool that embodies the current technical climate and accommodates varied learner needs. However, considerations must be made in the design, evaluation, and implementation of educational technology such that as a tool it is usable, accessible, appropriate, and promotes gains for all students as well as the larger educational system.

Universal design for learning (UDL) is a set of principles that address inequitable learning opportunities for educational technology. UDL theory, developed by the Center for Applied Special Technology (CAST), suggests that providing multiple means of representation, expression, and engagement allows greatest student access to learning opportunities (CAST, 2011). Cognitive tools, which embody these values, can be provided in a digital environment to increase opportunities for students who may struggle with static informational materials. However, while focus has been directed towards UDL tool development, and considerable gains have been accomplished with the establishment of UDL guidelines, the influence of the instructional design on the utilization of the tools remains largely unexplored.

With regards to the design of instructional material, cognitive load theory suggests that several cognitive mechanisms are involved in learning and understanding instructional material (Sweller, 1994). These cognitive processes allow information to move through the working memory, where current mental activity occurs, into the long-term memory, where information that has been successfully processed is stored (Carlson et al., 2003). In cognitive load theory, cognitive load is the demand placed on the working memory by instructional materials; as cognitive load increases, it is more challenging to manage information processing (Marcus, Cooper, & Sweller, 1996). In cognitive load theory there are three types of cognitive load (intrinsic, extraneous, and germane), which must be managed in order for information to successfully pass from the limited short term memory into the relatively unlimited long term memory and, consequently, for learning to occur. This successful passage of
information, from the short term memory to the long term memory supports the retention of learning.

The work presented here is part of a larger effort to design and evaluate digital educational material, specifically for teaching high school chemistry, that incorporates the needs of a wide spectrum of learning abilities. The present focus is on extending the findings from cognitive load research, which has largely been conducted with pencil and paper materials, to a digital format. The results are interpreted with respect to the derivation of design requirements for student-centered digital instructional material that utilizes UDL tools.

II. Cognitive Load Theory

When considering human factors in the design of instructional materials, as is in the case of educational technology, cognitive load theory (CLT) has many implications for the design of effective learning materials (Paas, Renkel, & Sweller, 2004). Derived from the cognitive processes that are needed for learning and understanding, CLT is based on properties of the short term and long term memory structures. Short term memory structures are described as limited capacity storage that can contain around 7 unrelated items (van Merrienboer & Ayres, 2005). Alternatively, long term memory structures are considered relatively unlimited and hold cognitive schemas, or knowledge structures, that vary in complexity and levels of automation.

Numerous studies have shown that instructional formats that ignore the considerations of these concepts of CLT and involve learners in cognitive activities often impose cognitive load in a manner that impedes learning (Leahy & Sweller, 2005; Mayer & Moreno, 2003; Paas et al., 2004; Sweller, 1999, 2003). Therefore, this problem should be amenable to design modifications that can improve performance outcomes in the tradition of human factors.

Within this framework, there are three primary types of cognitive load - intrinsic, extraneous, and germane. Intrinsic cognitive load, as described by Sweller (1994), is the inherent intellectual complexity of the instructions, which is dependent on the number of elements and interactivity of the elements in the instructional task. For example, the task of learning the location of the bones in the middle ear, which is comprised of the malleus, incus, and stapes has a low intrinsic load; the task of learning the bones in the wrist, of which there are eight, has a higher intrinsic load. Sweller further categorized cognitive load with an extraneous component, which is determined by the manner in which the information is presented, such as text or diagram. Lastly, Paas and Merrienboer (1994) found that by varying the type of instruction, a final type of cognitive load, germane, supports construction of knowledge structures by providing a means of motivation and increased engagement. This can be seen, for example, by providing a variety of problem situations in order to encourage construction of knowledge structures.

Each type of load has varying implications for the design of effective learning materials, and combined are considered additive. It should be noted, though, that both conditions of underload and overload are suspect to cause decreasing performance (Paas et al., 2004). Furthermore, any of the individual load components could independently produce a state of overload for the learner, as well as could the combined, total cognitive load. However, these levels can also be managed within the capacity of the short term memory. For example, in a high state of intrinsic load (complex content), a partially solved problem (extraneous load reduction) would remove the requirement of storing a stipulated number of elements in the working memory.

III. General Methodology

Using cognitive load theory as a conceptual foundation to explore UDL principles, the laboratory experiments were designed to further investigate the behavior of students performing different tasks that employ UDL tools. In order to do so, the research on CLT conducted by Carlson et al. (2003) on learning and understanding science instructional material was first replicated in the RoboBooks digital platform (see Software below). The intent of the experiment replication was to determine if intrinsic and extraneous load manipulations could be successfully conducted in the digital workbook in congruence with CLT as performed by Carlson. The replication, conducted in an initial laboratory session, included an extension into new content area, and the development of an additional task. The intent of the extension task was to determine if the principles of CLT were consistent across varying science tasks. In the first lab session, two designs, which have varying levels of extraneous load, were created to test three tasks; each with subtasks having low and high intrinsic load.

A second session was conducted to compare extraneous load manipulations conducted in the first session with an additional UDL component. In this session, two additional designs, in which UDL tools were employed, utilized the second and third tasks from the first session. These designs explored two UDL tools, highlight and concept map, under varying levels of intrinsic cognitive load. The objective of this session was to determine if the UDL tools in the second set of designs produced similar behavior when compared to the
varying levels of extraneous load in the first session designs.

Hardware
All lab experiments were conducted on 13” MacBook laptop computers with OS X Snow Leopard. An Apple Mouse was attached via the USB port. The experiment required both input from the external mouse to navigate the RoboBooks software and input from the built-in keyboard and web camera to respond to questions in the experiment.

Software
The RoboBooks software is driven by a LabView engine. The user interacts with an HTML front panel hosted within the Firefox web browser by a LabView server. In the web browser, the toolbar and navigation bars were disabled to reduce extraneous window features. Participants navigated through the software by using the mouse to access previous or successive pages, as indicated on the screen. During the test portions of the experiment and between subtasks, access to prior pages was disabled by removal of the “back” button.

Data Collection
Three primary dependent measures were collected to assess tasks within the four experiment conditions. Task time and test scores were used to quantify performance; a five-point subjective workload rating (SWR) was used to quantify mental workload, as perceived by the participant (T. Sheridan, personal communication, 2011). These dependent measures were used to compare the effect of the independent variables: intrinsic load (high and low conditions), extraneous load (high or low conditions) and UDL presentation (present or absent conditions).

Task Descriptions
The participants were asked to complete three tasks (molecular modeling, naming compounds, and chemical equations) successively in the RoboBooks environment. The first task was comprised of two subtasks, while Tasks 2 & 3 were comprised of four subtasks, in the order shown in Table 1. The first two subtasks were comprised of learning instructional material with low element interactivity (Subtask 1), which will forthwith be referred to as “Low I” to represent the low intrinsic load condition. Similarly learning instructional material with high element interactivity (Subtask 2) will be referred to as “High I” to represent the high intrinsic load condition. The Test subtasks, described as “open” or “closed” book tests, represent a first test where the instructional materials are allowed, and a subsequent test where instructional materials are not available.

<table>
<thead>
<tr>
<th>Task</th>
<th>Subtask 1 (Low I)</th>
<th>Subtask 2 (High I)</th>
<th>Subtask 3</th>
<th>Subtask 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1: Molecular Modeling</td>
<td>Model Simple Molecules</td>
<td>Model Complex Molecules</td>
<td>N/a</td>
<td>N/a</td>
</tr>
<tr>
<td>Task 2: Naming Compounds</td>
<td>Learn Prefixes</td>
<td>Learn Suffixes</td>
<td>Test 1: “Open Book”</td>
<td>Test 2: “Closed Book”</td>
</tr>
</tbody>
</table>

Preceding each primary task, the participants were provided a pre-task explanation page that explained the components of the task, in congruence with Carlson’s methodology. This explanation would describe the task in terms of subtasks and describe any particular components of the task that would otherwise not be included in the instructional material. This includes descriptions of the modeling pieces or definitions of key words that might be used in the instructions.

In each of 4 conditions, the same sequence of tasks, subtasks, and tests were presented to the participants. However, in Session 2, only Tasks 2 (Naming Compounds) and 3 (Chemical Equations) were required of the participants. In each design, all participants were provided the same pre-task instructions, subjective workload rating scales, and test questions. The instructional format for the Low I and High I subtasks, for each of the 3 tasks, varied with the objective for each of the 4 designs.

Test Scores
For Tests 1 & 2, scores were assigned per the metric defined by Carlson. Accordingly, 1 mark was assigned for each correct prefix and 1 mark for each correct suffix. A student could receive a total of 12 points for the prefix portion of Test 1, and 12 points for the suffix portion of Test 1. For Test 2, participants could receive a possible total of 6 points for the prefix and the suffix components of the test. Tests 3 & 4 were similarly constructed, such that the participant could independently receive up to 10 points for Test 3 and 5 points for Test 4 for identifying reactions and balancing equations. It is expected that test scores should be higher when intrinsic load is low and when extraneous load is reduced in high intrinsic conditions, when...
compared to low intrinsic load and low extraneous load, respectively.

Task Times

Task time was derived from the navigation tracking, which assigned a time stamp to each “page turn.” By taking the difference in time stamps, a page time was calculated for each of the participants. By summing the page times, a time per subtask was derived. It is expected that task times should be shorter when intrinsic load is low and when extraneous load is reduced in high intrinsic conditions, when compared to low intrinsic load and low extraneous load, respectively.

Subject Information

The participants were 20 students from Tufts University. Of the 20 subjects, 8 were male and 12 were female. Participant ages ranged from 19-35, and 6 of the participants were graduate students. The participants had varying amounts chemistry background and were categorized into novice and expert categories, accordingly, though performance means were aggregated. The number of experts and novices in each condition was balanced. 10 of the students were randomly assigned the “High E” Design condition and 10 were assigned the “Low E” Design condition.

IV. Results

The experimental design described above included an extension beyond Carlson’s initial findings. For the purposes of this paper, results from the first two subtasks (high I and low I) will be evaluated in the two conditions of low and high extraneous load.

The measures of interest in Task 1 were task time and subjective workload rating (SWR). (Note: as these subtasks were not followed by a subsequent content test, test score is not applicable for Task 1). The high extraneous load instructional format in the High E Design was compared to the low extraneous instructional format in Low E Design, as the between subject factors. Subtask 1 (Low I), the task with low intrinsic load, was compared to the high intrinsic load Subtask 2 (High I), as the within subject factors. Task time was compared in the between subject variable. SWR was compared in the between and within subject variables.

Table 2 shows mean seconds to task completion and average SWR (M) for each of the conditions, with associated standard deviations (SD). It should be noted that the direction of the means show a reduction in time to task and workload between High E and Low E Designs for both subtasks. Further comparisons show an increase in SWR between Low I and High I subtasks.

### Table 2. Means and standard deviations for task time and SWR in Task 1

<table>
<thead>
<tr>
<th>Group</th>
<th>Task Time Max</th>
<th>M</th>
<th>SD</th>
<th>SWR Max</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtask 1: Simple Model</td>
<td>551 340.7</td>
<td>103.9</td>
<td>263.0</td>
<td>293.1</td>
<td>89.9</td>
<td></td>
</tr>
<tr>
<td>SWR</td>
<td>7</td>
<td>2.5</td>
<td>0.7</td>
<td>1.7</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>Subtask 2: Complex Model</td>
<td>444 295.4</td>
<td>81.7</td>
<td>212.1</td>
<td>60.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWR</td>
<td>7</td>
<td>3.7</td>
<td>1.0</td>
<td>2.4</td>
<td>1.1</td>
<td></td>
</tr>
</tbody>
</table>

Note. n=10. Max = maximum.

Task time data was subjected to a 2 (instructional format [High E] vs. [Low E]) x 2 (Subtask 1 [Low I] vs. Subtask 2 [High I]) ANOVA, with repeated measures on the second factor. There was a significant main effect for the design, F(1, 18) = 293.1, MSE = 3086913.6, indicating that the diagram group (Low E) overall took less time than the text group (High E). (Note: the 0.05 level of significance is used throughout this paper). Results also showed a significant main effect for the subtask, F(1, 18) = 5.7, MSE = 23136.1, suggesting that overall Subtask 2 (High I) took longer to complete than Subtask 1 (Low I).

Subjective workload data was also analyzed in a 2 (instructional format [High E] vs. [Low E]) x 2 (Subtask 1 [Low I] vs. Subtask 2 [High I]) ANOVA. Results also showed a significant main effect for the design, F (1, 18) = 174.9, MSE = 255.0, indicating that overall the diagrammatic group (Low E) reported less mental workload than the text based group (High E). Results showed a significant main effect for the subtask, F(1, 18) = 33.5, MSE = 9.0, suggesting that overall subjects reported higher mental workload for Subtask 2 (High I). There was also a significant format x level of element interactivity interaction, F(1, 18) = 9.1, MSE = 13.2, suggesting that in high intrinsic load subtask (High I), the design of instructional material impacted the workload.

V. Conclusion

The findings for Task 1 indicate that diagrammatic instructional format, as presented in the Low E Design, resulted in the faster construction of high element interactive molecular modeling (High I). Overall, learners required less time to construct the high intrinsic load models presented in a diagrammatic form (Low E) and reported them to have lesser mental workload than the text based counterpart (High E). These findings were congruent with Carlson et al. (2003) and further support the following claims: (1) when compared to low element interactivity, high intrinsic load leads to decreased performance and higher subjective workload, (2) extraneous load reduction can promote performance gains and reduction in workload.
for tasks with high element interactivity. Notably, these findings were maintained when the media shifted from paper-and-pencil to digital workbook tasks.

Though mean differences were less distinguishable for Subtask 1, one finding by Carlson et al. (2003) was not supported in this experiment. There was no significant design x task interaction for task times, which Carlson supported the claim that greater improvement was found in the diagrammatic format (Low E) for Subtask 2 (High I) than for Subtask 1 (Low I). This may be due to a slightly smaller sample size in this experiment (n=10), as the means do trend in a similar manner to the Carlson data (n=12).

VI. Discussion

Deriving cognitive load components in learning tasks provides a means to systemically predict behavior and consequently devise improved opportunities for individualized learning. Moreover, by understanding UDL within the context of CLT, these opportunities can be extended to increase accessibility and best promote equitable learning opportunities. Ultimately, by approaching the design of digital instructional media from a cognitive load theory framework, active management of cognitive load is achieved through utilization of supplementary accessibility tools.

When considering the range of cognitive capacities, designing UDLs in digital instructional media within the context of CLT provides active mechanisms for students who may have limited capacity for information acquisition. For these students, management of cognitive load is increasingly imperative, especially when skills for schemata creation may also be limited. Yet, tools for managing cognitive load are additionally pertinent to aid all learners in managing cognitive load.

Moreover, a CLT approach has implications for the design of educational software from a human factors perspective for both curriculum and interface development. As shown in this research, learning tasks can be defined in terms of cognitive load components. The research findings suggest that tasks that are high in intrinsic load should be accompanied with extraneous load reduction; this includes utilizing multiple information channels, offloading memory demands, and the provision of cognitive aids. These componentized tasks and associated techniques for extraneous load reduction can be tested with users prior to production, in a true learner-centered design fashion.

References


